MULTIVALENT ELECTROSTATIC INTERACTIONS AND AUTOINHIBITION UNDERLIE PHASE SEPARATION OF NPM1 WITH R-RICH NUCLEOLAR PROTEIN SURF6

Authors:
Michael J. Wolek, Jaclyn A. Cika, and Richard W. Kriwacki

Faculty Sponsor:
Richard W. Kriwacki
Department of Structural Biology, St. Jude's Children's Research Hospital, Memphis, Tennessee

ABSTRACT

The nucleolus, the largest membrane-less organelle, is critical to cellular functioning as the site of ribosome biogenesis and as a fundamental cell stress sensor. It has previously been shown that the nucleolus assembles via liquid-liquid phase separation of its components from the surrounding nucleoplasm. Here, we show that two abundant nucleolar proteins often upregulated in certain cancers, Nucleophosmin (NPM1) and Surfeit locus protein 6 (Surf6), interact via multivalent electrostatic interactions and form phase-separated droplets**.** Utilizing microscopic techniques, we propose a novel interaction of the C-terminal nucleic acid binding domain of NPM1 to modulate self valency through autoinhibition of its acidic residues.

INTRODUCTION

A hallmark of the Eukaryotic cell is the division of cellular processes into membrane-bound organelles. Robust compartmentalization in cells is vital for efficient spatial and temporal regulation of cellular biochemistry. For instance, the physical barrier of the nuclear membrane serves as a mechanism to regulate the pool of mRNA transcripts available for translation, thus regulating cellular protein levels [1]. Additionally, the endomembrane system utilizes a series of membranes for intra- and extra-cellular protein localization. In addition to organelles that rely on lipid barriers to centralize their components and biochemical reactions, there is also a class of structures called membrane-less organelles which do not rely on lipid membranes to contain macromolecular components [2].

Membrane-less organelles, often referred to as Ribonucleoprotein (RNP) bodies, are fluid-like congregations of proteins and nucleic acids found throughout the cytoplasm (e.g. P bodies, stress granules, and germ granules) and nucleoplasm (e.g. nucleoli, Cajal bodies, nuclear speckles, and PML bodies) [2, 3]. Although our understanding of the physical properties and assembly mechanisms of these structures is limited, membrane-less organelles are generally thought to be important in biological systems for facilitating reaction efficiencies due to their ability to concentrate reaction components [4]. Despite lacking a lipid bilayer to separate their contents from the surrounding milieu, RNP bodies have also been shown to be important in facilitating signaling in response to cellular stress caused by changes in internal equilibrium. In the event of cellular stress, proteins or RNAs can be released from or sequestered within RNP bodies, altering the concentration of macromolecules in the cellular milieu, thus creating a signal to initiate a stress response [2]. Although all play different cellular roles, alterations of the composition or self-assembly of these RNP bodies has been linked to debilitating diseases including Huntington's, spinal muscular dystrophy, and neurodegenerative disorders including amyotrophic lateral sclerosis (ALS) [5-7].

The nucleolus is the largest of the membrane-less organelles and serves as the center for ribosome biogenesis, as well as a cell stress sensor [2]. The nucleolus is organized into three distinct subcompartments: the dense fibrillar component (DFC) surrounds the fibrillar center (FC), both of which are surrounded by the granular component (GC). This subcompartmentalization is thought to facilitate

ribosome biogenesis by concentrating rRNA modifying and processing machinery, allowing for rRNA transcription and processing to occur in a vectoral manner extending from the FC out to the GC [3]. Ribosome biogenesis begins with a pre-rRNA transcript at the border between the FC and DFC, which is then spliced and modified to a smaller subset of rRNAs in the DFC, which are further assembled and folded with ribosomal proteins into pre-ribosomal particles in the GC [3]. Since ribosome production is critical for protein production and rapid cell growth, pathologists use nucleolar size, shape, and quantity as indicators for cancer growth and malignancy [4]. Furthermore, certain ribosomal proteins and the 5S rRNA are linked to tumorigenesis as they are important for modulating p53 anti-tumor protein activity. It is presumed that p53-dependent stress sensing mechanisms are incorporated in the nucleolus to halt the energetically taxing process of ribosome production during unfavorable growth conditions [2]. Furthermore, disruption of nucleolar function to impair both ribosome biogenesis and/or the cell stress response can lead to ribosomopathies such as Diamond-Blackfan anemia [8].

In order to understand the molecular basis underlying these disorders we need to understand the mechanism behind membrane-less organelle assembly and disassembly. An increasing body of evidence suggests that liquid-liquid phase separation is the general mechanism governing membrane-less organelle assembly. Liquid-liquid phase separation (LLPS) is a physical phenomenon that describes two or more co-existing, demixed liquid phases, similar to the separation of oil and vinegar in vinaigrette. This separation of protein solutions into a protein-rich phase, demixed from the surrounding medium, occurs when a critical concentration or temperature threshold is crossed, whereupon self-solvation of the protein is thermodynamically preferred over its solvation in the surrounding buffer [2, 9]. The critical concentration for LLPS can be influenced by a number of factors, including temperature, ionic strength, and macromolecule valency [2]. Specific for individual proteins, the phase separation boundary is the varying critical value at which phase separation is observed, taking into consideration protein concentration, ionic strength, and protein valency [2]. Multivalent interactions have been shown to facilitate heterogeneous LLPS, wherein both components are required to exhibit multivalency [7, 10, 11] Furthermore, across macromolecules, certain characteristic features have been identified to promote LLPS including electrostatic and multivalent interactions, nucleic acid binding domains (particularly RNA), and proteins with stretches of low-complexity amino acid sequence [3, 7, 10-12]. For example, nucleolar localization signals (NoLS), short, low-complexity stretches of mainly arginines and lysines, have been shown to be necessary for localization of certain proteins to the nucleolus [13]. Despite growing evidence that phase separation governs assembly and protein localization to nucleoli, more information is needed to understand the molecular mechanisms underlying this process. In order to advance toward therapeutic applications targeting hyperactive nucleoli in various malignancies, a basic structural and mechanistic understanding of the phase-separated nucleolus is necessary.

A multifunctional protein, NPM1, has been shown to localize to the GC of the nucleolus and play key roles in nucleolar assembly and function [11, 14]. NPM1 participates in p53-dependent and independent cell stress responses, genome stability, regulation of centrosome duplication, tumor suppression, and nucleolar structure [15, 16]. Deletion of NPM1 has been associated with disruption of the nucleolar morphology and is embryonic lethal in mice, suggesting a structural significance in nucleoli [16]. Likely due to its multifarious roles, *NPM1* is commonly mutated in various hematopoietic and solid cancers, and is the most common gene mutated in adult acute myeloid leukemia (AML) [11]. A large proportion of mutations are located in the C-terminal nucleic-acid binding domain, resulting in NBD unfolding as well as aberrant localization of NPM1 to the cytoplasm, a common feature in adult AML cells [17]. In breast cancer in particular, cancerous cells often contain NPM1 overexpression, which has been shown to drastically stimulate c-MYC-induced hyperproliferation [15]. However, other studies suggest NPM1's oncogenic role depends on a complex system of conditions, at times playing a tumor suppressive role in cells [18].

The structural features of NPM1 are important when considering its dual role in nucleolar integrity and cellular stress responses. NPM1 contains three major domains: An N-terminal oligomerization domain (OD), a central intrinsically disordered region (IDR), and a C-terminal nucleic

acid binding domain (NBD). NPM1 also contains three acidic tracts: acidic tract 1 (A1) is located in the OD whereas A2 and A3 are within the IDR (Figure 1). NPM1's multivalent character is further amplified due to its ability to form homopentamers via its OD [11]. Recently, peptides containing arginine-rich motifs (R-motifs) were shown to bind to a truncated NPM1 construct through A1 and A2 [11]. Consistent with NPM1's functional roles, many of its binding partners are known players in ribosome biogenesis, as well as in tumor suppression [11]. Particularly with respect to its role in ribosome biogenesis, NPM1 has been shown to have a higher affinity for single-stranded RNA in addition to containing endogenous endoribonuclease activity, preferentially cleaving pre-ribosomal RNA [17, 19].

Over 100 binding partners of NPM1 have been identified through a cellular pull-down assay coupled with mass spectrometry analysis, most of which contain multivalent R-motifs, a structural feature that mediates LLPS with NPM1 [11]. Surf6 is a highly basic nucleolar protein containing sequence-level and structural features of a protein important for nucleolar assembly and function [20, 21]. Although its specific functional role is debated, Surf6 has been shown to localize to the GC during interphase and follow the path of major rRNA processing proteins, such as NPM1 during mitosis [13]. Studies of the Surf6 homolog in *Saccharomyces cerevisiae* suggest that it is is involved in the 35S pre-rRNA synthesis and large ribosomal subunit assembly [20]. Furthermore, Surf6 and NPM1, UBF (the main RNA pol I cofactor), and EBP2 (rRNA processing factor) have been shown to co-IP, further supporting the role of Surf6 in ribosome biogenesis [14]. Interestingly, *Surf6* is overexpressed in embryonic cells as well as in human tumor cells, which require increased ribosome production, suggesting Surf6 plays an important role in ribosome biogenesis [21].

The Surf6 protein is characterized as largely disordered, containing multiple lysine and arginine-rich motifs [21]. The combination of high disorder propensity as well as several R-rich interaction motifs resemble characteristics of proteins prone to undergo LLPS [7, 11]. On the other hand, NPM1 exhibits regions of disorder as well as multivalency, however NPM1's multivalent character is unique due to the presence of multiple acidic residues. Considering these inherent structural features in NPM1 and Surf6, the R-motifs of Surf6 have the potential to interact and phase separate with the multiple acidic residues on NPM1, providing a useful model for nucleolar GC phase separation [3, 7, 11, 12].

Recently, NPM1 has been shown to interact with R-rich peptides derived from nucleolar binding partners to form distinct, phase separated droplets *in vitro* [9]. Moreover, the OD of NPM1 is required for phase separation with proteins and rRNA and the NDB is required for phase separation with rRNA [3, 11]. Multivalent and electrostatic interactions between the acidic tracts A2 and A3 of NPM1's IDR and R-motif containing peptides has been shown to drive phase separation *in vitro*. Since Surf6 is a binding partner of NPM1 and localizes to the GC, it might contribute to the formation of the GC's liquid structure, however the mechanism by which these molecular interactions with NPM1 govern the fluidity of the GC are not fully understood [11].

Our goal, therefore, is to gain further understanding into the molecular mechanism of LLPS of NPM1 with the R-rich nucleolar protein, Surf6. Through further understanding, we hope to gain insight into not only the fluidity of the GC, but also the fundamental structural behavior by which NPM1 derives its diverse functional roles, especially when related to cancer.

RESULTS

NPM1 C-terminal NBD Modulates Phase Separation and Intermolecular Interactions in NPM1/Surf6$^{1-183}$ Droplets

The C-terminal NBD of NPM1 has been previously shown to be critical for phase separation with rRNA. Here we investigated its structural significance–if any–with respect to droplet composition and dynamics with a Surf6 construct consisting of amino acid residues 1-183, which will be referred to as Surf6$^{1-183}$. We already know R-rich proteins bind to the proximal, N-terminal acidic domains of NPM1. Because the C-terminal domain is positively charged to promote interactions with rRNA and it is located at the end of a highly disordered stretch of amino acids, the C-terminal NBD might interfere with protein partners binding to the distal acidic residue A3 in NPM1. Surf6$^{1-183}$ is predicted to be highly disordered (Fig. 1B)

and contains several R-rich motifs which have previously been shown to interact with the acidic residues A1 and A2 of NPM1 [11]. To gain insight into the role of the NBD and acidic tract A3 in protein dynamics within droplets, we performed a series of *in vitro* phase separation assays comparing the physical properties of NPM1/Surf6$^{1\text{-}183}$ phase separated droplets using two NPM1 constructs: N294 (the full-length protein) and N240 (a truncated version lacking the 54 residue C-terminal domain) (Figure 1).

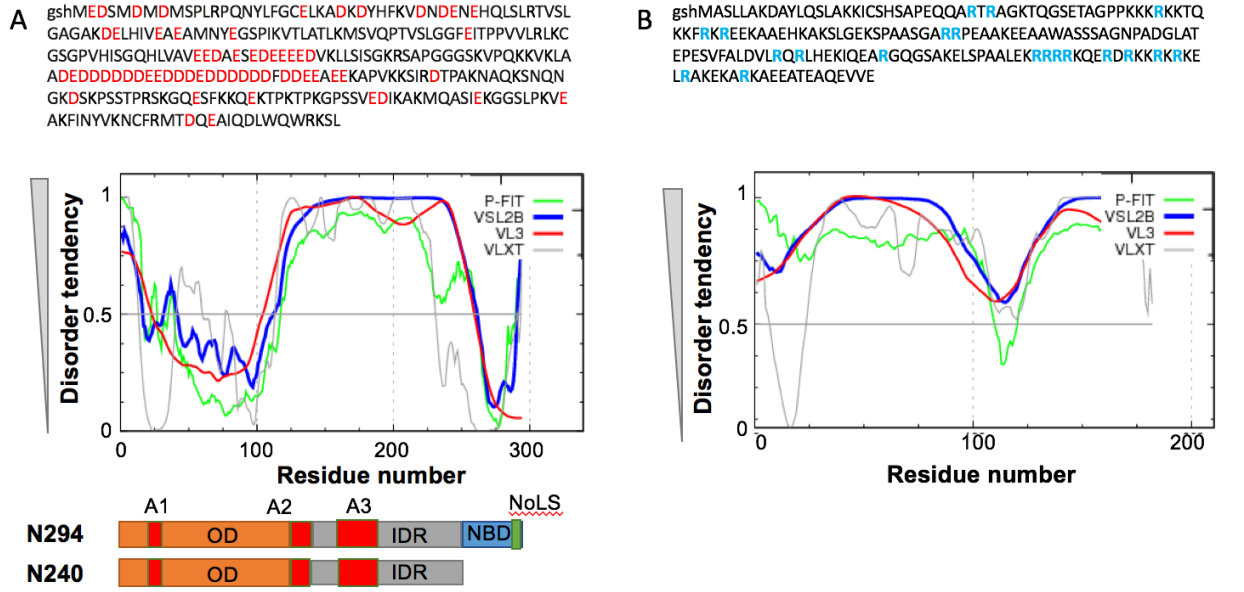


Figure 1. Structural details of NPM1 and Surf6$^{1\text{-}183}$

(A) Sequence of N294 with acidic residues shown in red (top). Disorder plot (middle) shows the predicted disorder tendency along the protein, using various algorithms: P-FIT (green), VSL2B (blue), VL3 (red), and VLXT (grey) from http://www.disprot.org/metapredictor.php. Schematic diagram of monomeric NPM1 constructs (bottom).

(B) Sequence of Surf6$^{1\text{-}183}$ construct with arginine residues shown in blue (top). Disorder plot shows predicted disorder tendency along the protein, using the same algorithms used in fig. 1A (middle).

NPM1 NBD is Important in Determining Valency of NPM1 when Interacting with Surf6$^{1\text{-}183}$

To see if the C-terminal domain contributes to NPM1/Surf6$^{1\text{-}183}$ interaction, we analyzed the ability to phase separate and the physical properties of phase-separated protein droplets using confocal microscopy at fixed concentrations of NPM1 while titrating in an R-rich protein partner, Surf6$^{1\text{-}183}$. No observable difference between the NPM1 constructs was expected, as the NBD of NPM1 is only known to play a role in phase separation with RNA. However, compared to droplets formed with N294 and Surf6$^{1\text{-}183}$, N240-containing droplets fail to exhibit robust phase separation when Surf6$^{1\text{-}183}$ is in stoichiometric deficiency (Figure 2). This observation is indicative of a shift in the phase separation boundary for NPM1 *vs.* N240. Since protein levels, salt concentration, and all other conditions remained constant comparing the NPM1 to Surf6$^{1\text{-}183}$ ratio of 1:0.5, we propose that the valency and number of binding sites for R-motifs in the two NPM1 constructs are different, resulting in the observed shift in the phase boundary.

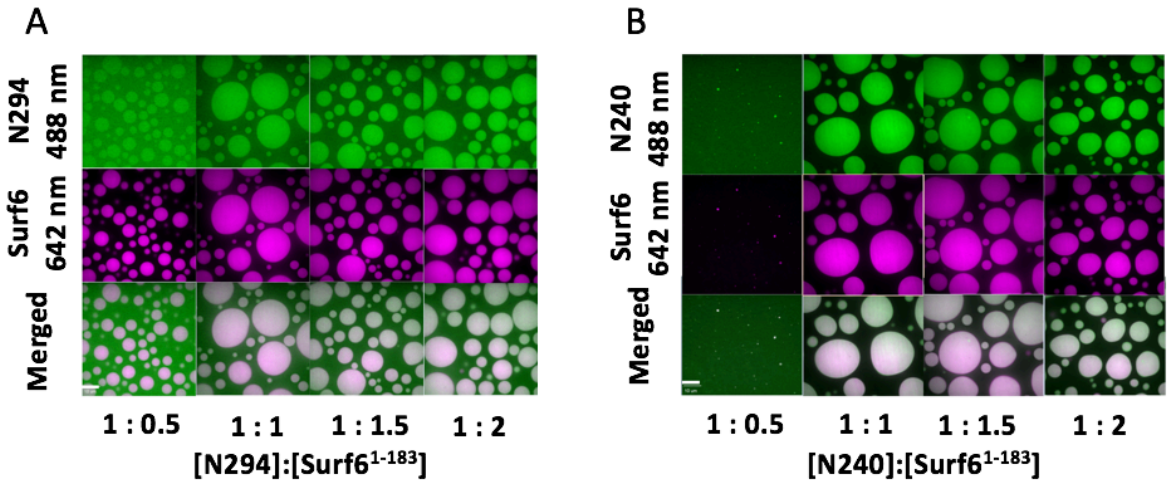


Figure 2. The NBD of NPM1 is Important for Modulating Phase Separation with R-rich Surf6$^{1-183}$
In vitro images of N294 and Surf6$^{1-183}$ mixtures (left) and N240 and Surf6$^{1-183}$ mixtures (right) at varying concentrations of Surf6$^{1-183}$ at 20μM NPM1, imaged for NPM1 (Alexa Fluor 488), Surf6$^{1-183}$ (Alexa Fluor 647), and combined NPM1/ Surf6$^{1-183}$ fluorescence. Scale bar, 10 μm.

The Removal of NPM1's NBD Facilitates Increased Interaction with Surf6$^{1-183}$
Further to understand what might be different between phase-separated droplets formed between the NPM1 constructs, we examined the kinetic properties of the proteins encapsulated within these droplets using additional microscopy techniques. Fluorescence recovery after photobleaching (FRAP) is a common method of measuring protein dynamics in which a region of interest containing fluorescently labeled molecules is selectively photobleached with a high intensity laser. Over time, the recovery of non-bleached molecules into the region of interest is measured, from which one can obtain the mobile fraction of protein as well as the apparent diffusion rate of the freely diffusing protein [22]. As such, FRAP recovery curves were generated to determine whether differences in molecular kinetics exist between N294 and N240 when both NPM1 constructs phase separate with Surf6$^{1-183}$.

FRAP experiments of these droplets *in vitro* show significant differences in droplet physical properties between NPM1 constructs. When the recovery curves between N294 and N240-containing droplets are compared, there is only a subtle reduction in the percent recovery of N240 droplets compared to N294 (Figure 3A). However, the apparent rate of diffusion and the mobile fraction of NPM1 inside droplets was reduced an average of 31% and 12%, respectively, in N240 droplets versus N294 (Figure 3B). These results suggest that the freely diffusing N240 particles are larger in size since diffusion is inversely proportional to size. Additionally, when the entire droplet was photobleached, there was reduced percent and time of recovery in N240 droplets compared to N294 droplets, suggesting differences in diffusion at the interface between the buffer and droplet (Figure 3C).

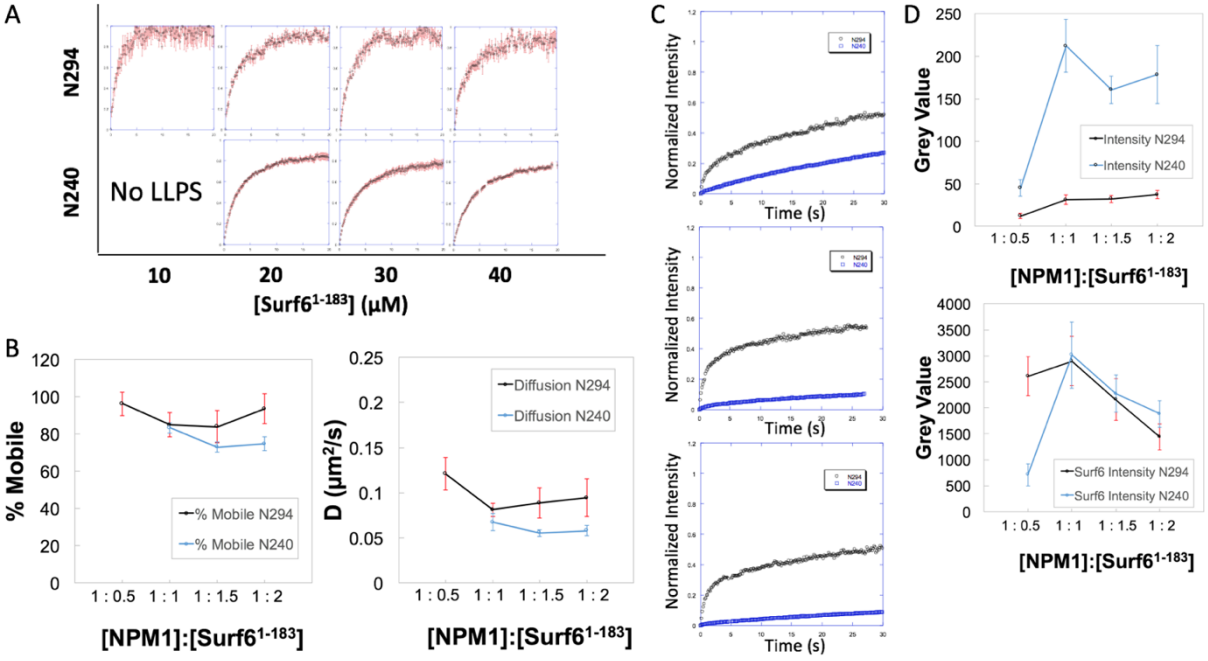
Altered droplet structure extends beyond the diffusing fraction of macromolecules in the N240 droplets compared to full-length N294 droplets. Cross-sections within droplets were analyzed for both NPM1 and Surf6$^{1\text{-}183}$ intensities to determine if differences in droplet composition varied between the two NPM1 constructs. Interestingly, there appears to be a higher ratio of NPM1 construct to Surf6$^{1\text{-}183}$ in droplets containing the C-terminal truncation, N240, when compared to full-length N294 droplets. This result suggests that more Surf6$^{1\text{-}183}$ binding sites are accessible to N240 compared to N294 (Figure 3D).

Figure 3. N240 Droplets Exhibit Different protein Dynamics Compared to N294

(A) Average FRAP recovery curves for 20 µM N294 (top) and N240 (bottom) 30 minutes after mixing samples at various concentrations of Surf6$^{1\text{-}183}$. Small region of interest within the droplets was photobleached and monitored over the duration of the experiment. N=5.

(B) Average % mobile fraction of NPM1 (left) and apparent rate of diffusion (tau) of NPM1 (right) extrapolated from the FRAP recovery data in (A) as a function of Surf6$^{1\text{-}183}$ concentration. % Mobile fraction and diffusion of NPM1 compared between N294 (black) and N240 (blue). N240 + Surf6$^{1\text{-}183}$ droplets did not phase separate into robust droplet structures at 10 µM Surf6$^{1\text{-}183}$. Error bars represent SD. N=10-15 per point.

(C) Entire droplet FRAP recovery curves for NPM1 at 20 µM Surf6$^{1\text{-}183}$ (top), 30 µM Surf6$^{1\text{-}183}$ (middle), and 40 µM Surf6$^{1\text{-}183}$ (bottom). N294 droplet recovery (black) compared to N240 droplet recovery (blue) at varying concentrations of Surf6$^{1\text{-}183}$.



(D) Relative intensity of NPM1 (top) and Surf6$^{1\text{-}183}$ (bottom) in droplets at varying concentrations of Surf6$^{1\text{-}183}$. Intensities compared between N294 (black) and N240 (blue) conditions. Error bars represent SD. N=5.

NPM1 C-terminal NBD reduces electrostatic interactions in the absence of rRNA

Next, we investigated the significance of electrostatic interactions on the intermolecular interactions mediating droplet structure and assembly. The negatively charged R-motifs of Surf6$^{1\text{-}183}$ are expected to interact with NPM1's acidic residues, facilitating LLPS electrostatically. However, if the NBD of NPM1 is also interacting with A3 electrostatically, N240 should exhibit properties suggesting it has more sites available for electrostatic interactions. To investigate this hypothesis, a series of salt titration experiments

were performed *in vitro* comparing N294 and N240-containing droplets. To ensure that hydrophobic interactions were not being promoted due to an increase in buffer ionic strength, hexanediol was added to the buffer, which interferes with weak hydrophobic interactions. N294 and N240 droplets were observed using a salt gradient starting at 75 mM and increasing at an interval of 25 mM per sample up to 200 mM (above physiological salt conditions). By increasing the salt concentration, the electrostatic interactions between molecules should be gradually inhibited. Thus, by measuring the component protein concentrations in droplets with increasing salt concentrations, we can compare the strength of electrostatic interactions of N294 with N240. Droplets were viewed and imaged using confocal microscopy after 30 minutes post-mixing. Images were analyzed and the average fluorescence intensity for each droplet component were extracted across all salt conditions, in droplets formed with each of both NPM1 constructs.

Across both NPM1 constructs, as salt was increased and electrostatic interactions between NPM1 and Surf6$^{1\text{-}183}$ were disrupted, both Surf6$^{1\text{-}183}$ and NPM1 concentrations reach saturating levels, then decrease rapidly above a certain critical concentration as droplet dissociation is promoted and the phase boundary is traversed (Figure 4A & B). Additionally, at salt concentrations not nearing the edge of the phase boundary (75-150 mM NaCl) N240 levels inside droplets were 10-20 fold higher than N294 droplets, while Surf6 levels remained relatively consistent between the two conditions (Figure 4C & D). Since N240 is predicted to have more binding sites available for multivalent interactions, it makes sense that NaCl exhibits an increased effect on N240 droplets, as shown by smaller droplets compared to N294 at 175 and 200 mM NaCl as well as a steeper decline in NPM1 levels in N240 droplets (Figure 4B & C).

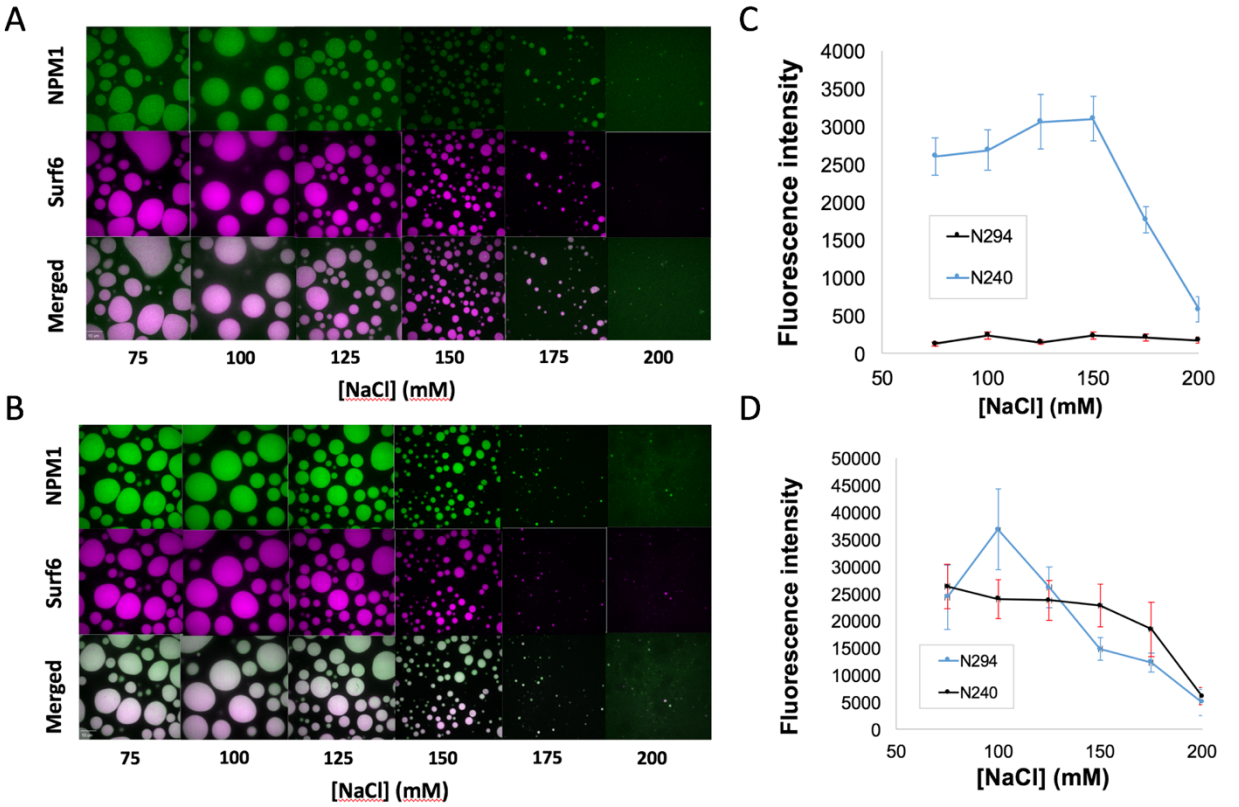


Figure 4. Salt Affects N240 Droplet Dynamics More Harshly, Showing Altered Electrostatic Interactions Compared to N294

(A-B) Salt titrations of N294/ Surf6$^{1\text{-}183}$ droplets (A) and N240/ Surf6$^{1\text{-}183}$ droplets (B) at a constant concentration of 20 µM NPM1 and 20 µM Surf6$^{1\text{-}183}$ were imaged for NPM1, Surf6$^{1\text{-}183}$, and combined

NPM1/ Surf6$^{1-183}$ fluorescence. Scale bar, 10 µm. Images of *in vitro* droplets in buffer 10 mM Tris, 150 mM NaCl, 4 mM DTT, 5% hexanediol, pH 7.5.
(C-D) Relative intensities of NPM1 (C) and Surf6$^{1-183}$ (D) at increasing concentration of NaCl identical to that of (A & B). Intensities compared between N294 (black) and N240 (blue) conditions. N=5.

Full-length NPM1 phase separates in the presence of both Surf6$^{1-183}$ and rRNA

As rRNA is a major nucleolar component in nucleoli, we investigated the significance of the interplay between NPM1, Surf6, and rRNA together. Since rRNA has been shown to interact with the NBD of NPM1, we expect rRNA to be integrated into phase-separated droplets along with Surf6$^{1-183}$. To test this, we performed confocal microscopy of droplets containing the combinations of Surf6$^{1-183}$ and rRNA alone compared to Surf6$^{1-183}$, rRNA, and N294 together. As expected, both binary and ternary mixtures yielded droplets (Figure 5). However, the ternary mixture droplets were smaller in size than those formed without rRNA. Interestingly, Surf6$^{1-183}$ phase separates with rRNA in the absence of N294. (Figure 5). This makes sense given Surf6 and rRNA are ultimately phase separated in the GC, and may simply be the result of rRNA interacting with the positive R-rich motifs of Surf6$^{1-183}$.

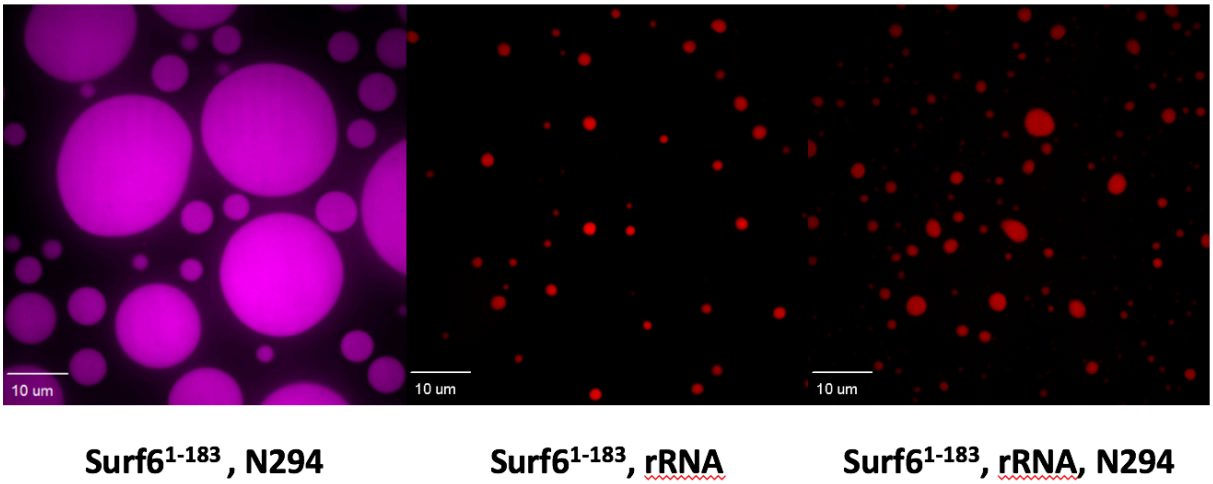


Figure 5. N294 Facilitates Surf6$^{1-183}$ and rRNA into Localized Liquid-Liquid Phase Separated Droplets

In vitro images of 20 µM Surf6$^{1-183}$ mixed with 20 µM N294 (left), 100 µg/mL rRNA (middle), and 100 µg/mL rRNA with 20 µM N294 (right). Scale bar, 10 µm. Surf6 fluorescence was monitored. Buffer solution of 10 mM Tris, 150 mM NaCl, 4 mM DTT, pH 7.5.

DISCUSSION

Unlike many other proteins shown to be critical in membrane-less organelle assembly and function, NPM1 is able to interact with two different classes of molecules found in the nucleolus through multivalent, electrostatic interactions. The acidic multivalent nature of NPM1 is critical for interacting with R-rich, positively charged nucleolar proteins, whereas its positively charged nucleic acid binding domain plays a role in interacting with rRNA. Because NPM1 contains both acidic stretches necessary to interact with its binding partners, as well as a C-terminal disordered stretch and structured NBD, NPM1 may also have the ability to interact with itself via these regions.

In vitro experiments varying both Surf6$^{1-183}$ concentration and salt levels support the idea that the NBD modulates the valency of NPM1 in the absence of rRNA. When Surf6$^{1-183}$ concentration was varied, the phase boundary of NPM1/Surf6$^{1-183}$ droplets shifted with N240 compared to N294. However, the C-terminal domain of N294 should not be interacting with Surf6$^{1-183}$ as both are largely basic in nature and should repel one another. And yet, the reduced rate of diffusion of N240 in droplets and the increased

sensitivity of N240 to salt suggests N240 pentamers are larger structures and have increased valency compared to N294, resulting in more interactions with Surf6$^{1-183}$ and thus reduced ability to freely diffuse within droplets. These results support the hypothesis that NPM1's C-terminal domain folds back on itself, covering up entirely or partially certain acidic tracts in an autoinhibitory manner. Since the NBD is presumed to cover or partially cover a patch of acidic residues, exposing that patch in N240 creates the ability to form more extensive cross-linking networks between adjacent pentamers when compared to N294. The differences between protein dynamics in the absence of rRNA once the C-terminal domain and NPM1's presumed ability to interact with itself is removed suggests that NPM1 has adopted the ability to regulate the extent of interactions between itself and R-rich proteins.

Whether or not this mechanism is a way to regulate rRNA binding has yet to be determined. One possible way to explore the A3/NBD interaction and its role in modulating rRNA binding would be to use an NPM1 construct with a mutated A3, eliminating its ability to interact with itself via these regions. Future experiments will be performed to address if incorporation of an NPM1 C-terminal domain binding partner has the ability to pull back this domain and free-up covered acidic residues, thus allowing for more R-proteins to bind. Additionally, it would be interesting to observe the effects of NPM1 C-terminal truncation in cells, as this may mimic the effects of the aberrant cytoplasmic localization on NPM1 as seen in adult AML cells. As both Surf6 and NPM1 are nucleolar components, it would make sense if Surf6 is also capable of phase separating with rRNA. Perhaps the ability of both Surf6 and NPM1 to bind to rRNA serves as a mechanism for traversing rRNA trough the GC and out of the nucleolus via a hand-off of rRNA when Surf6 binds to NPM1.

MATERIALS AND METHODS

Cloning, Protein Expression, Purification, and Fluorescent Labeling

NPM1 constructs were previously subcloned in the pET28 vector (Novagen) with an N-terminal poly-His tag, however the PreScission Protease cleavage site was replaced with a TEV Protease cleavage site [11]. NPM1 was expressed in BL21(DE3) *E. coli*. Protein expression was induced at OD_{600} ~ 0.6 with 100 mg/L IPTG (Goldbio) and cultures were incubated overnight at 18^{o}C. Bacterial pellets were harvested by centrifugation and lysed by sonication in 10 mM Tris, 300 mM NaCl, 5 mM ß-mercapto-ethanol (BME), pH 7.5. N294 was purified by Ni-NTA affinity chromatography, eluting with 500 mM Imidazole containing lysis buffer. Eluted proteins were dialyzed overnight against 10 mM Tris, 150 mM NaCl, 2 mM DTT pH 7.5 in the presence of TEV protease to remove the poly-His tag. N294 was further purified by HPLC with a C4 column (Waters). The protein was lyophilized, resuspended in 6M Guanidine HCl, and dialyzed against 10 mM Tris, 150 mM NaCl, 2 mM DTT, pH 7.5. Proper refolding was checked by CD spectroscopy. N294 C21/275T (previously made following the same protocol as N294) was labeled with Alexa Fluor488 C5 maleimide on Cys104 according to the manufacturer's protocol (Thermo Fisher Scientific) in 10 mM Tris, 150 mM NaCl, pH 7.5 buffer to maintain the folded conformation of NPM1. Labeling efficiency was determined by UV-Vis spectroscopy. To ensure that approximately 1 monomer per pentamer was labeled with Alexa dye, labeled protein solution with 13% labeled were unfolded in 6M GdnHCl and refolded as mentioned above.

Surf6$^{1-183}$ construct was subcloned in the pET28a vector with a TEV cleavage site between NdeI and SalI. Surf6$^{1-183}$ was expressed in Rosetta(DE3) *E. coli* and protein expression was induced at OD_{600} ~ 0.6 with 100 mg/L IPTG (Goldbio). Cultures were incubated overnight at 18^{o}C. Bacterial pellets were harvested by centrifugation and resuspended in 25 mM NaP, 0.1% Triton X-100 and incubated at room temperature for 30 minutes. Protein was pelleted by centrifugation and resuspended in 25 mM NaP, 300 mM NaCl, 6 M Guanidine HCl, 5 mM BME, pH 7.0 and lysed by sonication. Surf6$^{1-183}$ was purified by Ni-NTA affinity chromatography, eluting with 500 mM Imidazole containing lysis buffer. Eluted proteins were dialyzed overnight against 10 mM Na phosphate, 150 mM NaCl, 2 mM DTT pH 7.5 in the presence of TEV protease to remove the poly-His tag. Surf6$^{1-183}$ was further purified using HPLC with a C4 column (Waters) and lyophilized overnight. Surf6$^{1-183}$ was then resuspended in 6 M Guanidine HCl and dialyzed against 25 mM Na Phosphate, 300 mM NaCl, 2 mM DTT, pH 7.0. Surf6$^{1-183}$ was N-terminally labeled with

Alexa Fluor647 NHS-ester following the manufacture's protocol (ThermoFisher). Labeling efficiency was determined by UV-Vis spectroscopy. Left two images in Figure 5 used A555 labeled Surf6$^{1-183}$, fluorescing red, whereas the rightmost image used A647 labeled Surf6$^{1-183}$ and fluoresced pink.

Slide Coating for Microscopy

CultureWell 16-well chambered coverglass (Grace Biolabs), coated with PlusOne Repel-Silane ES (GE Healthcare) were used. Slides were first filled with detergent and heated in a microwave oven for ~1 minute. Slides were then rinsed with MilliQ water, followed by several rinses with 70% ethanol, followed by another wash with MilliQ water. Next, slides were dried with nitrogen gas and the wells were filled approximately half-way with PlusOne Repel-Silane ES. After sitting at room temperature for ~2 minutes, the silane was discarded and slides were rinsed with 3 washes of 70% ethanol, 3 washes of MilliQ water, and one wash of 70% ethanol. The wells were then dried with nitrogen gas and filled with 1% PF-127. Slides were capped and sealed with parafilm around the edges until use.

Light Microscopy of *in vitro* droplets

Confocal images used for phase separation assays were collected on a Mariannas 1 microscope equipped with a CSU-X spinning disk using a 100X objective with oil. 75 µL samples were incubated at room temperature for 30 minutes prior to analysis. Images were analyzed using SlideBook6 software.

FRAP

FRAP experiments were performed using the same instrument mentioned above. For region of interest FRAP (ROI), spots of radius ~0.5 µm were bleached with a 488 laser for A488 labeled N294 & N240 and a 561 and a 642 laser for A555 & A647 labeled Surf6$^{1-183}$, respectively in SlideBook 6. Entire droplet FRAP was done using the same respective lasers, but at varying radii dependent on the size of the droplets. Data was exported from SlideBook6 and imported into Excel and KaleidaGraph for further processing. FRAP curves were background subtracted and corrected for autophotobleaching of the sample. Data points were normalized to the initial fluorescence intensity before bleaching. Relative intensity was fit to a single exponential curve $F=A(1-e^{-t/\tau})$ where tau is the apparent recovery time and A is the maximum recovery [22]. The apparent diffusion coefficient was estimated as $D_{app}=r^2/\tau$ [22]. Relative protein levels within droplets was determined by taking the average relative intensity values from the cross section of the inside of a droplet in SlideBook 6 normalized to background intensities.

ACKNOWLEDGEMENTS

Diana Mitrea provided helpful comments on manuscript.

REFERENCES

[1] Terry LJ, Shows EB, Wente SR. Crossing the nuclear envelope: hierarchical regulation of nucleocytoplasmic transport. Science. 2007;318:1412-6.
[2] Mitrea DM, Kriwacki RW. Phase separation in biology; functional organization of a higher order. Cell Commun Signal. 2016;14:1.
[3] Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, et al. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell. 2016;165:1686-97.
[4] Brangwynne CP, Mitchison TJ, Hyman AA. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc Natl Acad Sci U S A. 2011;108:4334-9.
[5] Savas JN, Makusky A, Ottosen S, Baillat D, Then F, Krainc D, et al. Huntington's disease protein contributes to RNA-mediated gene silencing through association with Argonaute and P bodies. Proc Natl Acad Sci U S A. 2008;105:10820-5.
[6] Gall JG. Cajal bodies: the first 100 years. Annu Rev Cell Dev Biol. 2000;16:273-300.
[7] Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell. 2015;163:123-33.
[8] Ebert B, Lipton JM. Diamond Blackfan anemia and ribosome biogenesis: introduction. Semin Hematol. 2011;48:73-4.
[9] Hyman AA, Weber CA, Julicher F. Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol. 2014;30:39-58.
[10] Li P, Banjade S, Cheng HC, Kim S, Chen B, Guo L, et al. Phase transitions in the assembly of multivalent signalling proteins. Nature. 2012;483:336-40.
[11] Mitrea DM, Cika JA, Guy CS, Ban D, Banerjee PR, Stanley CB, et al. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. Elife. 2016;5.
[12] Zhang H, Elbaum-Garfinkle S, Langdon EM, Taylor N, Occhipinti P, Bridges AA, et al. RNA Controls PolyQ Protein Phase Transitions. Mol Cell. 2015;60:220-30.
[13] Scott MS, Boisvert FM, McDowall MD, Lamond AI, Barton GJ. Characterization and prediction of protein nucleolar localization sequences. Nucleic Acids Res. 2010;38:7388-99.
[14] Kordyukova MY, Polzikov MA, Shishova KV, Zatsepina OV. Functional significance of the human nucleolar protein SURF6, the key member of the SURF6 protein family in eukaryotes. Dokl Biochem Biophys. 2014;455:65-7.
[15] Li Z, Boone D, Hann SR. Nucleophosmin interacts directly with c-Myc and controls c-Myc-induced hyperproliferation and transformation. Proc Natl Acad Sci U S A. 2008;105:18794-9.
[16] Mitrea DM, Grace CR, Buljan M, Yun MK, Pytel NJ, Satumba J, et al. Structural polymorphism in the N-terminal oligomerization domain of NPM1. Proc Natl Acad Sci U S A. 2014;111:4466-71.
[17] Gallo A, Lo Sterzo C, Mori M, Di Matteo A, Bertini I, Banci L, et al. Structure of nucleophosmin DNA-binding domain and analysis of its complex with a G-quadruplex sequence from the c-MYC promoter. J Biol Chem. 2012;287:26539-48.
[18] Karhemo PR, Rivinoja A, Lundin J, Hyvonen M, Chernenko A, Lammi J, et al. An extensive tumor array analysis supports tumor suppressive role for nucleophosmin in breast cancer. Am J Pathol. 2011;179:1004-14.
[19] Savkur RS, Olson MO. Preferential cleavage in pre-ribosomal RNA byprotein B23 endoribonuclease. Nucleic Acids Res. 1998;26:4508-15.
[20] Polzikov MA, Kordyukova MY, Zavalishina LE, Magoulas C, Zatsepina OV. Development of novel mouse hybridomas producing monoclonal antibodies specific to human and mouse nucleolar protein SURF-6. Hybridoma (Larchmt). 2012;31:48-53.
[21] Polzikov M, Magoulas C, Zatsepina O. The nucleolar protein SURF-6 is essential for viability in mouse NIH/3T3 cells. Mol Biol Rep. 2007;34:155-60.
[22] Lippincott-Schwartz J, Altan-Bonnet N, Patterson GH. Photobleaching and photoactivation: following protein dynamics in living cells. Nat Cell Biol. 2003;Suppl:S7-14.